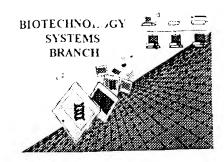
0570

RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/901, 484Source: 01/66Date Processed by STIC: 10/23/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216. PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax) PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER</u> <u>VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-modernt. Checker all no publications to about a common betings in Commuter Readable form

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION SERIAL NUMBER: 09/90/484
ATTN: NEW RULES CASES	S: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAL
lWrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5Variable Length	Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6PatentIn 2.0 "bug"	A "bug" in Patentin version 2.0 has caused fire <220>-<223> acction to be missing from amino acid sequences(s)
7Skipped Sequences (OLD RULES)	Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped
	Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8Skipped Sequences (NEW RULES)	Sequence(s) missing. If Intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
0Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
1Use of <220>	Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
Patentin 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
3Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

OIPE

RAW SEQUENCE LISTING

DATE: 10/23/2001 TIME: 16:31:00

PATENT APPLICATION: US/09/901,484

Input Set : A:\GEN-T1111XC3D2-seqlist.txt Output Set: N:\CRF3\10232001\I901484.raw

3 <110> APPLICANT: Cohen, Daniel Blumenfeld, Marta Ilya, Chumakov Bougueleret, Lydie 8 <120> TITLE OF INVENTION: PROSTATE CANCER GENE 10 <130> FILE REFERENCE: GEN-T111XC3D2 12 <140> CURRENT APPLICATION NUMBER: 09/901,484 13 <141> CURRENT FILING DATE: 2001-07-09 15 <150> PRIOR APPLICATION NUMBER: 08/996,306 16 <151> PRIOR FILING DATE: 1997-12-22 18 <150> PRIOR APPLICATION NUMBER: 60/099,658 19 <151> PRIOR FILING DATE: 1998-09-09 21 <150> PRIOR APPLICATION NUMBER: 09/218,207 22 <151> PRIOR FILING DATE: 1998-12-22 24 <150> PRIOR APPLICATION NUMBER: 09/338,907 25 <151> PRIOR FILING DATE: 1999-06-23 27 <150> PRIOR APPLICATION NUMBER: 09/853,526 28 <151> PRIOR FILING DATE: 2001-05-11 30 <160> NUMBER OF SEQ ID NOS: 578 32 <170> SOFTWARE: Patent.pm

ERRORED SEQUENCES

2650 <210> SEQ ID NO: 69 2651 <211> LENGTH: 5226 2652 <212> TYPE: DNA 2653 <213> ORGANISM: Homo sapiens 2655 <400> SEQUENCE: 69 2656 atgetateed tagtaeteed caequactee ata ego tae eta eta eec ago ate 54 2657 Met Arg Tyr Leu Leu Pro Ser Val 2658 2659 gtg etc etg gge acg geg dec acc tac gtg ttg gec tgg ggg gtc tgg 102 2660 Val Leu Leu Gly Thr Ala Pro Thr Tyr Val Leu Ala Trp Gly Val Trp 266115 2662 egg etg etc tec gec tte etg ecc gec ege tte tac caa geg etg gac 150 2663 Arg Leu Leu Ser Ala Phe Leu Pro Ala Arg Phe Tyr Gln Ala Leu Asp 30 35 2665 gad egg oto tac tgc gtd tac dag agd atg gtg oto ttd ttd ttd gag 198 2666 Asp Arg Leu Tyr Cys Val Tyr Gln Ser Met Val Leu Phe Phe Phe Glu 2667 4.5 50 the contract of the contract o

ak la liya Kala Adh lan ling liya ling ayaa ka Ala Rah dha li la land liddi kaal liga kiya 2673 Ж'n 80

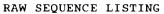
RAW SEQUENCE LISTING

DATE: 10/23/2001 PATENT APPLICATION: US/09/901,484 TIME: 16:31:01

Input Set : A:\GEN-T111XC3D2-seqlist.txt Output Set: N:\CRF3\10232001\1901484.raw

		-		gac Asp													342
.1676		90		-			95					100		-			
2677	cgc	tac	gtg	ctg	aaa	gaa	ggg	tta	aaa	tgg	ctg	cca	ttg	tat	ggg	tgt	391)
2678	Arg	Tyr	Val	Leu	Lys	Glu	Gly	Leu	Lys	Trp	Leu	Pro	Leu	Tyr	Gly	Cys	
2679	105					110					115					120	
2680	tac	ttt	gct	cag	cat	gga	gga	atc	tat	gta	aag	cgc	agt	gcc	aaa	ttt	438
2681	Гуr	Phe	Ala	Gln	His	Gly	Gly	Ile	Tyr	Val	Lys	Arg	Ser	Ala	Lys	Phe	
2682					125					130					135		
				gag													486
2684	Asn	Glu	Lys	Glu	Met	Arg	Asn	Lys	Leu	Gln	Ser	Tyr	Val	Asp	Ala	Gly	
J685				140					145					150			
2686	act	cca	atg	tat	ctt	gtg	att	ttt	cca	gaa	ggt	aca	agg	tat	aat	cca	5.34
2687	Thr	Pro	Met	Tyr	Leu	Val	Ile	Phe	Pro	Glu	Gly	Thr	Arg	Tyr	Asn	Pro	
2688			155					160					165				
2689	gag	caa	aca	aaa	gtc	ctt	tca	gct	agt	cag	gca	ttt	gct	gcc	caa	cgt	582
2690	Glu	Gln	Thr	Lys	Val	Leu	Ser	Ala	Ser	Gln	Ala	Phe	Ala	Ala	Gln	Arg	
2691		170					175					180					
				gta													630
	_	Leu	Ala	Val	Leu	_	His	Val	Leu	Thr		Arg	Ile	Lys	Ala		
2694						190					195					200	
				ttt													678
	His	Val	Ala	Phe	_	Cys	Met	Lys	Asn	=	Leu	Asp	Ala	Ile	_	Asp	
2697					205					210					215		
				gtt									-	cgaa	agaga	agt	727
	Val	Thr	Val	Val	Tyr	Glu	Gly	Lys	_	Asp	Gly	Gly	*				
2700				220					225								5 0.0
		_					-			-						tgatc	787
	,			-	, ,				-	,			_	, ,	-	atgaac	847
	,	_			,	-	-	_	•		_	_		,	_	aagaa	907
	-	_		_	, , ,	_					_		_	_		accat	967
			-				-	-	_	-		_				ggaage	1027
	_	-									-					taaag	1087
		-	_		-				-							gegget	$\frac{1147}{1207}$
					_		_				-					gttgat	1207
		-														aacat	$\frac{136}{1327}$
		-														igttte	1387
											_					tgtgt	130.
			,			,							, -			ittttc ittagt	1447 1507
	_						-	_		-					.,	iaaatt	1567
									-				_		-	laaall laaaag	$\frac{1567}{1627}$
																tatat	1627
		_			-											cacaca	1747
																acaca	1807
		-	_		-		_									.aycay :t.tqtc	1867
- / I 🐣	adel	. и.са с	a Cail	COULC	ayu	با با	yad	чась	. Lad	гусса	icyc	all	-ayto	4 C - L	こいいしんし	. 4. 4. 4. 4. 4.	1007

DATE: 10/23/2001



PATENT APPLICATION: US/09/901,484 TIME: 16:31:01

Input Set : A:\GEN-T1111XC3D2-seqlist.txt
Output Set: N:\CRF3\10232001\1901484.raw

2723	tcacagaata	ttcattcaga	agtegegttt	ctgtagtgtg	gtggattccc	actgggctct	2107
2724	ggtccttccc	ttggatcccg	teagtggtge	tgctcagcgg	cttgcacgta	gacttgctag	2167
2725	gaagaaatgo	agagecagec	tgtgctgccc	actttcagag	ttgaactctt	taagcccttg	2227
2726	tgagtgggct	tcaccagcta	ctgcagaggc	attttgcatt	tgtctgtgtc	aagaagttca	2287
2727	ccttctcaag	ccagtgaaat	acagacttaa	ttcgtcatga	ctgaacgaat	ttgtttattt	2347
2728	cccattaggt	ttagtggagc	tacacattaa	tatgtatcgc	cttagagcaa	gagctgtgtt	1407
2729	ccaggaacca	gatcacgatt	tttagccatg	gaacaatata	teccatggga	gaagacettt	2467
2730	cagtgtgaac	tgttctattt	ttgtgttata	atttaaactt	cgatttcctc	atagtccttt	2527
	aagttgacat	_			_	_	2587
	cattaaacat	-			-		2647
	attaagtgac			_			2707
	tctttgaaga		•			-	2767
	atggacacte					-	2827
	tecteatget			2 2 2 2	222	-	2887
	totttqttaa			, ,	_	_	2947
	actttgagag	2 2 2 2		-		-	3007
2739	tgacctgcat	aagatcactt	gaatgttagg	tttcatagaa	ctatactaat	cttctcacaa	3067
	aaggtetata	-		-			3127
	agetteteet	_		_	-	-	31.87
	tacctacctg						3247
	acagettagt	-			-	_	3307
2744	ttcagaaaag	ggtgtgtttg	gatgaaagta	aaaaaaaaa	taaaatcttt	cactgtctct	3367
	aatggctgtg					-	3427
2746	aaaataggct	taatgactgg	ccctgcattc	ttcacaatat	ttttccctaa	gctttgagca	3487
2747	aagttttaaa	aaaatacact	aaaataatca	aaactgttaa	gcagtatatt	agtttggtta	3547
	tataaattca						3607
2749	taatcattaa	gtgatctcag	tgaaacatgt	caaatgcctt	aaattaacta	agttggtgaa	3667
2750	taaaagtgcc	gatctggcta	actcttacac	catacatact	gatagttttt	catatgtttc	3727
2751	atttccatgt	gatttttaaa	atttagagtg	gcaacaattt	tgcttaatat	gggttacata	3787
2752	agctttattt	tttcctttgt	tcataattat	attctttgaa	taggtctgtg	tcaatcaagt	3847
2753	gatctaacta	gactgatcat	agatagaagg	aaataaggcc	aagttcaaga	ccagcctggg	3907
2754	caacatatcg	agaacctgtc	tacaaaaaaa	ttaaaaaaaa	ttagccaggc	atggtggcgt	3967
2755	acactgagta	gtttgtccca	gctactcggg	agggtgaggt	gggaggatcg	cttcagccca	4027
2756	ggaggttgag	attgcagtga	gccatggaca	taccactgca	ctacageeta	ggtaacagca	4087
2757	cgagacccca	actcttagaa	aatgaaaagg	aaatatagaa	atataaaatt	tgcttattat	4147
2758	agacacacag	t.aactcccag	atatgtacca	caaaaaatgt	gaaaagagag	agaaatgtct	4207
	accaaagcag						4267
2760	tttgttttta	gtagtgttta	gattgaagat	tgagtgaaat	attttcttgg	cagatattcc	4327
2761	gtatctggtg	gaaagctaca	atgcaatgtc	gttgtagttt	tgcatggctt	gctttataaa	4387
2762	caagattttt	tctccctcct	tttgggccag	ttttcattac	gagtaactca	cactttttga	4447
2763	ttaaagaact	tgaaattacg	ttatcactta	gtataattga	cattatatag	agactatgta	4507
2764	acatgcaatc	attagaatca	aaattagtac	tttggtcaaa	atatttacaa	cattcacata	4567
2765	cttgtcaaat	attcatgtaa	ttaactgaat	ttaaaacctt	caactattat	gaagtgctcg	4627
2766	tctgtacaat	cgctaattta	ctcagtttag	agtagctaca	actcttcgat	actatcatca	4687
	atatttgaca				,		4747
2768	gaqtcatata	tgaggtcaaa	gacatatacc	ttgttattat	aatatgtata	ctataataat	4807
			1.0	7.7			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/901,484

DATE: 10/23/2001 TIME: 16:31:01

Input Set : A:\GEN-T111XC3D2-seqlist.txt Output Set: N:\CRF3\10232001\I901484.raw

2772 tttatatgtg aatatgtaag atatgttetg caattttata aatgtteatg tetttttta 5047 2773 aaaaaqgtgc tatcgaaatt etgtgtetee ageaggeaag aataettgae taactetttt 5107 2774 tgtetettta tggtatttte agaataaagt etgaettgtg tttttgagat tattggtgee 5167 E--> 2775 tcattaattc agcaataaag gaaaatatgc atttcaaaaa naaaaaaaa aaaaaaaaa All item 9 on Eng Sunnay Sheet

 $16099 < 210 \cdot SEQ ID NO: 578$

16100 -: 211 · LENGTH: 19

16101 - 2212 - TYPE: DNA

16102 <2213 > ORGANISM: Homo Sapiens

16104 - 220 FEATURE:

16105 <221> NAME/KEY: misc_feature

16106 + 2225 LOCATION: 1...19

16107 <223> OTHER INFORMATION: potential microsequencing oligo for 99-222-109.mis2

in the second of the second

16109 -: 400> SEQUENCE: 578

16110 ctgaagaaat tcatatcgt

delite

19

E--> 16112/283

10/23/01

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/901,484

DATE: 10/23/2001 TIME: 16:31:04

Input Set : A:\GEN-T111XC3D2-seqlist.txt
Output Set: N:\CRF3\10232001\1901484.raw

L:453 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 L:454 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 L:476 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 L:805 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 L:1102 M:283 W: Missing Blank Line separator, <400> field identifier L:1254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 L:1323 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:4 L:1419 M:283 W: Missing Blank Line separator, <400> field identifier L:1737 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 L:1971 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 L:2162 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:39 L:2775 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:69 L:2782 M:283 W: Missing Blank Line separator, <400> field identifier L:2819 M:283 W: Missing Blank Line separator, <400> field identifier L:2908 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:72 $L:2911\ M:341\ W:\ (46)\ "n"\ or\ "Xaa"\ used, for SEQ\ ID#:72$ L:2964 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:2965 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:2985 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3038 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 $L\!:\!3052$ M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3053 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3054 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3056 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3085 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3095 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 $L\!:\!3096$ M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3110 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3128 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3179 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 L:3228 M:283 W: Missing Blank Line separator, <400> field identifier L:3630 M:283 W: Missing Blank Line separator, +400> field identifier L:3641 M:283 W. Missing Blank Line separator, <400> field identifier L:3650 M·283 W. M:ssing Blank Line separator, <400 field identifier L:3659 M:283 W: Missing Blank Line separator, <400 - field identifier L:3668 M 283 W Missing Blank Line separator, <400° field identifier L:3677 M:283 W Missing Blank Line separator, <400> field identifier L:3687 M 283 W: Missing Blank Line separator, <400 - field identifier L:3811 M:283 W: Missing Blank Line separator, <4400> field identifier L:3820 M:283 W: Missing Blank Line separator, <400> field identifier L:3829 M:283 W: Missing Blank Line separator, $\pm 400 >$ field identifier L:3838 M:283 W: Missing Blank Line separator, <400> field identifier L:5482 M.283 W: Missing Blank Line separator, 4400: field identifier L:5648 M·283 W: Missing Blank Line separator, <400: field identifier

the state of the s

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/901,484

DATE: 10/23/2001 TIME: 16:31:04

Input Set : A:\GEN-T111XC3D2-seqlist.txt
Output Set: N:\CRF3\10232001\I901484.raw

L:7828 M:283 W: Missing Blank Line separator, <400> field identifier L:7869 M:283 W: Missing Blank Line separator, <400> field identifier L:7965 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 L:7984 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 L:8038 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8053 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8074 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8075 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 $L\!:\!8179$ M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8232 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8313 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8325 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 $L\!:\!8660$ M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 L:8741 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:184 L:16112 M:254 E: No. of Bases conflict, LENGTH:Input:283 Counted:19 SEQ:578